SAGE and Maximum Entropy Algorithmic Comparison for Adaptive Antenna Array Calibration

Alberto Antón[(1)], Ramón Martínez[(2)], Alberto Torre[(1)], Isabel García-Rojo[(1)], Rafael Hernández[(1)], Iván Albarrán[(1)]
[(1)] Ingeniería y Servicios Aeroespaciales, S.A. (INSA), Madrid, Spain
alberto.anton@insa.org
[(2)] Radiation Group, Universidad Politécnica de Madrid (UPM), Madrid, Spain
ramon@gr.ssr.upm.es

Abstract— *Adaptive antenna arrays, also known as smart antennae, are becoming progressively a viable alternative to larger antenna structures in a wide number of different scenarios. However, in order to take full advantage of the arrays' interference rejection and SNR (Signal-to-Noise Ratio) improvement capabilities, that is, in order to perform a correct signal beamforming via complex weights' manipulation, the antenna structure needs a precise calibration procedure. In this paper, two algorithms, based on efficient MLE (Maximum Likelihood Estimation) and MEE (Maximum Entropy Estimation) are analyzed and simulated, in order to obtain their relative performance.*

1. INTRODUCTION

Since their appearance at the beginning of the sixties, satellites communications systems have experimented an increasing demand and application scope. In the near future, such systems will require higher binary rates and, therefore, link antennae with higher gain and performance. Besides, the expected proliferation of new LEO (Low Earth Orbit) satellite constellations will make mandatory the use of very flexible ground stations with interference cancellation capabilities.

At the present time, satellite tracking stations take advantage of large reflector antennae. These pose a number of impairments regarding their mechanical complexity, lower flexibility and network efficiency, and higher cost when certain diameter limits are surpassed [1].

The reflector cost has an exponential dependence with its diameter, derived in part from the need of a complex pointing mechanic procedure in order to track the orbiting satellite. The maintenance of such a system can become very expensive. Furthermore, surface errors on the reflector during the design process increase with the size of the antenna, which limits its operation to high frequency bands. Finally, reflector antennae can track only one satellite at a time, so the efficiency of the earth segment is reduced.

Alternative antenna technologies shall thus be considered for a near future application. A possibility makes use of antenna arrays with smaller individual radiating elements, combined with digital signal processing. When implemented algorithms make use of signal statistics and a priori information to adapt the system's behavior, it is known as adaptive arraying.

Adaptive arrays with reduced-size radiating elements are becoming a viable alternative in several applications. Known as smart antennae in wireless communications, they seem especially well suited for that fast-varying multiuser cellular scenario, and constitute the founding stone for the MIMO (Multiple Input Multiple Output) systems and the increased channel capacity they theoretically promise. On the other hand, some RADAR transceivers make use of several thousand antennae, frequently optimized following stochastic techniques and subarray processing, in order to obtain real-time and multiple targets detection, acquisition and tracking. In radio-astronomy, distributed arrays of small parabolic antennae are used for interferometric purposes.

Two are the main problems solved by antenna arrays. First and foremost is their ability for radiation pattern beamforming and spatially white noise filtering. If the array factor maximum is pointed towards the desired signal DoA (Direction of Arrival) an optimal SNR gain is reached. The second problem is the input signals' DoA estimation. Classical beamscan algorithms have a spatial resolution limited by the Rayleigh criterion. However, algorithms based on signal correlation subspaces such as the well-known MUSIC (MUltiple SIgnal Classification) outperform that limitation and offer attractive super-resolution features.

The previously alluded advantages though require a perfect knowledge of the input signal model, which depends on statistical, electromagnetic, physical and geometrical parameters. In real systems many errors sources are present. Unfortunately, both spatial beamforming and DoA

978-1-4244-5128-9/10/$26.00 ©2010 IEEE

estimation algorithms severely degrade their performance, even if only small errors are present. Three possible solutions can be considered:

- To replace the spatial beamforming algorithms by time or blind reference ones. The first can still cancel uncorrelated interferences, but generally require a pre-demodulation not easily performed on low SNR scenarios. The latter only work without interferences. Anyhow, this solution does not fit in the DoA estimation problem.
- To implement robust algorithms. However, the most complex and efficient versions need large arrays with much redundancy. Besides, some robust DoA estimation algorithms additionally require some a priori information about the array model.
- To calibrate the array. This is the solution that will be considered in this paper, since it is the most complete and, frequently, the only one that can be practically implemented.

Numerous array calibration algorithms are scattered throughout the scientific literature. In this paper, we will center our attention on two of them, the SAGE (Space-Alternating Generalized Expectation-maximization) and the MEC (Maximum Entropy Calibration) algorithms. The paper is thus organized. Section 2 presents some revision about the calibration problem and what causes it. Section 3 analyzes the SAGE algorithm, its origins and its application to adaptive arrays. Section 4 describes MEC principles. In section 5 simulation results are showed. In section 6 conclusions and future works are drawn. Finally, a brief appendix explores some CRLB (Cramér-Rao Lower Bound) considerations.

2. THE CALIBRATION PROBLEM

Firstly, the array signal model is presented. Taking bold letters to describe vectors and matrices, the input signal in an N-element array can be modelled as:

$$\boldsymbol{x}[k] = f[k] \times \boldsymbol{So} + \sum_{l=1}^{L} i_l[k] \times \boldsymbol{Si}_l + \boldsymbol{n}[k] \qquad (1)$$

$f[k]$ is the sampled desired signal, with an explicit discrete-time dependence, $i_l[k]$ is the l-th interference signal, $\mathbf{n}[k]$ is generally a white Gaussian noise column vector, of size N, with statistically independent components, and $\mathbf{So}$ and $\mathbf{Si}_l$ are the column SV (Steering Vectors) of the desired and interference signals, with their i-th vector component defined as:

$$(\boldsymbol{SV})_i = \exp\left\{ j \frac{2\pi}{\lambda} (sin(\vartheta_0)\, cos(\varphi_0)\, \Delta x_i + sin(\vartheta_0)\, sin(\varphi_0)\, \Delta y_i + \cdots \right.$$
$$\left. \ldots + cos(\vartheta_0)\, \Delta z_i) \right\} \qquad (2)$$

λ is the signal wavelength, implying a narrowband model, ϑ_0 and φ_0 the corresponding signal DoA in spherical angles, and finally Δx_i, Δy_i and Δz_i are the position differences in Cartesian coordinates between the i-th array element and the user-defined reference one.

Spatial beamforming and DoA estimation algorithms use spatial a priori information of the array's electromagnetic environment, synthesized in the **SV** [2], [3]. The correct construction of these vectors implies perfect knowledge of the array's electrical and geometrical properties. Since this is usually not the case in real systems, errors arise:

- Radiating elements' geometrical positions' errors.
- Antennae DoA dependent's phase centers.
- Gain and phase offsets in array branches, I/Q modulation imbalance.
- Antennae coupling.
- Signal wavefront distortion, multipath.
- Quantization errors.
- Etc.

More information can be found in [4]. Trying to estimate and compensate some or all of these errors is known as the array calibration problem. This paper will be mainly concerned about position errors.

3. SAGE CALIBRATION ALGORITHM

Array calibration is a problem of parameter estimation. The most popular solution is the MLE technique, or, in a more general frame, with random unknown parameters, the MAP (Maximum A Posteriori) algorithm. In fact, it can be shown that MLE is a particular case of MAP, where no a priori information of the parameters PDF (Probability Density Function) is available. Thus, if MAP is used, careful notice must be taken concerning statistical assumptions. Of course, hybrid methods can be considered in multi-parametric problems. Likewise, certain estimations allow ignoring parasitic parameters, the values of which need not to be known. The estimators performance, concerning standard deviation or root mean square error, depending on whether MLE or MAP are considered, is lower-bounded by the CRLB when unbiased estimations can be derived. In that sense, it may be said that, for a particular problem, *if* an optimal unbiased parameter estimator exists, MLE/MAP provide it and they are limited by the CRLB [2].

MLE starts with the optimization of the log-likelihood function which, for complex zero-mean Gaussian random processes, is written as:

$$\widehat{\boldsymbol{\theta}} = argmax_{\boldsymbol{\theta}}[L(\boldsymbol{\theta})] = argmax_{\boldsymbol{\theta}}\{-[\ln\det(\boldsymbol{K}(\boldsymbol{\theta}) +$$
$$\left. \tfrac{1}{K}\textstyle\sum_{k=1}^{K} \boldsymbol{x}_k{}^H \boldsymbol{K}^{-1}(\boldsymbol{\theta})\boldsymbol{x}_k \right)]\} \qquad (3)$$

$\boldsymbol{\theta}$ is the unknown parameters' vector and $\mathbf{K}$ the process $\mathbf{x}$'s autocovariance matrix. Complex Gaussian statistics are well suited for arrays' applications. If MAP is considered, the last equation is written as:

$$\widehat{\boldsymbol{\theta}} = argmax_{\boldsymbol{\theta}}[L(\boldsymbol{\theta})] = argmax_{\boldsymbol{\theta}}\{-[\ln\det(\boldsymbol{K}(\boldsymbol{\theta}) +$$
$$\left. \tfrac{1}{K}\textstyle\sum_{k=1}^{K} \boldsymbol{x}_k{}^H \boldsymbol{K}^{-1}(\boldsymbol{\theta})\boldsymbol{x}_k \right)] + \ln p_{\boldsymbol{\theta}}(\boldsymbol{\theta})\} \qquad (4)$$

More information regarding its application to array calibration can be found in reference [5].

These methods are very powerful, though their main drawback is their computing complexity, due to the usual high dimensionality of the problems at hand. In order to cope with this, efficient implementations of MLE/MAP non-linear optimization were sought. One of such was the EM (Expectation Maximization) algorithm [6]. The idea is to decouple the original multi-parametric optimization into simpler sub-optimizations with reduced dimensionality. Decoupling is performed from the available, or "incomplete", data snapshots $\boldsymbol{x}_k$, which are transformed into "complete" data $\boldsymbol{y}_{ki}$, where i is the sub-optimization index. This transformation is not unique, and therefore must be designed for each particular problem. The technique reduces to maximizing the conditioned likelihood function of $\boldsymbol{y}_k$ given observed data $\boldsymbol{x}_k$ and current parameter estimation $\widehat{\boldsymbol{\theta}}^{(n)}$:

$$\widehat{\boldsymbol{\theta}}^{(n+1)} = argmax_{\boldsymbol{\theta}}\, E\left[\ln p_{\boldsymbol{y}_k}\left(\boldsymbol{y}_k : \boldsymbol{\theta} | \boldsymbol{x}_k : \widehat{\boldsymbol{\theta}}^{(n)}\right)\right] \quad (5)$$

Where $E[\ ]$ is the ensemble average or expectation operator.

SAGE brings the EM idea one step further. While EM decouples parameters but estimates them simultaneously, SAGE defines several processing cycles where only one subgroup is optimized. The advantage behind this is that each subgroup can be associated with a different signal space, called "hidden or augmented space" [7], clearly desirable when different kinds of parameters must be obtained. It can be proved that SAGE preserves the stability and computational complexity of EM, while it can significantly increase its convergence speed in certain scenarios [8].

In reference [9] SAGE is used for array position calibration, although it could be applied as a global calibration algorithm defining appropriate cycles. Calibration is performed using beacon signals from unknown DoAs. Therefore, two cycles are defined: beacons' DoAs and antenna positions' estimation cycles. More on beacon calibration theory can be found in [10].

4. MEC ALGORITHM

Entropy notion related with telecommunications was firstly introduced by Shannon in 1948, as a valid method to measure the uncertainty associated with a random variable with a given PDF. If a PDF is chosen to model a process, more entropy implies less a priori suppositions are being taken. Alternatively, random states with low entropy give more information if they show up, because their related uncertainty is lower. These concepts have been applied to the spectral estimation of random processes, for example in the classical periodogram method, which estimates autocorrelation with sample averages:

$$\hat{R}(k) = \frac{1}{K-k}\sum_{i=1}^{K-k}(\boldsymbol{x})_i\,(\boldsymbol{x}^H)_{i+k} \quad (6)$$

Its main drawback is higher imprecision for high values of lag index "k", since less samples are used. Windowing techniques can reduce this problem, although spectral resolution is lost. Bearing this in mind, Burg suggested an alternative. Instead of making high k-valued autocorrelations zero, he let them be equal to the coefficients which maximized the process entropy, i.e., those random processes which made less a priori assumptions, given the imposed low k-valued autocorrelation restrictions [11]. The obtained process is called autoregressive or Gauss-Markov, and satisfies the Yule-Walker equations:

$$\hat{R}(0) = -\sum_{k=1}^{N-1} a_k R(-k) + \sigma^2 \quad (7)$$

$$\hat{R}(l) = -\sum_{k=1}^{N-1} a_k R(l-k), \quad l = 1, \dots, N-1 \quad (8)$$

As can be seen, it is an N-unknowns N equations system: a_1, a_2, … , a_{N-1} and σ^2. In matrix form [12]:

$$\mathbf{R}\begin{pmatrix} 1 \\ a_1 \\ \vdots \\ a_{N-1} \end{pmatrix} = \begin{pmatrix} \sigma^2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (9)$$

Following the Wiener-Khintchine theorem, the process spectrum is:

$$S(\omega) = \frac{\sigma^2}{\left|1+\sum_{k=1}^{N-1} a_k e^{-jk\omega}\right|^2}, \quad -\pi \le \omega \le \pi \quad (10)$$

This problem has been applied in arrays to make DoA estimates with the spatial spectrum S(θ) [13]. As a novelty, MEC algorithm, instead of obtaining a spectrum as a function of DoAs, estimates one as a function of the considered calibration space, for example antennae positions:

$$S(\boldsymbol{p}) = S\left(\left[p_x\ p_y\ p_z\right]^T\right) = \frac{\sigma^2}{\left|1+a_1 e^{-j\boldsymbol{k}_o^T\boldsymbol{p}}\right|^2} \quad (11)$$

$\boldsymbol{p}$ is the Cartesian positions' vector, and $\boldsymbol{k}_o$ is the wavenumber vector, defined as:

$$\boldsymbol{k}_o = \frac{2\pi}{\lambda}[sin(\vartheta_0)\,cos(\varphi_0)\ \ sin(\vartheta_0)\,sin(\varphi_0)\ \ cos(\vartheta_0)]^T \quad (12)$$

Spectrum optimization is performed decoupling each antenna, so that only one a pair is considered each time. Next figure shows this spectrum for a 10-element ULA (Uniform Linear Array) and 0 dB of SNR. A block size of 1024 samples of a Gaussian white process, three beacon signals and uniform position errors with ±0.45d of amplitude were used, where d is the half-wavelength inter-element distance.

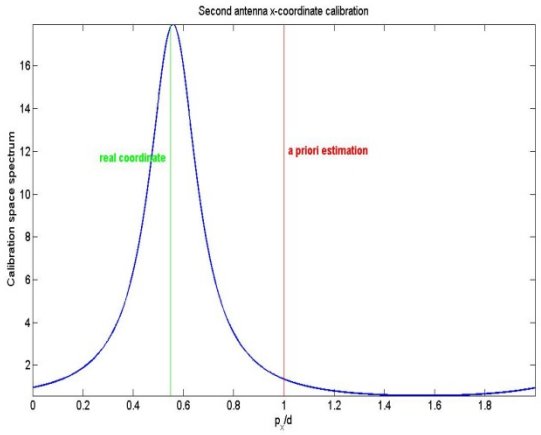

Figure 1 – Maximum entropy spectrum for sensor calibration

5. SIMULATIONS

Simulations have been carried out in Matlab, for a 10-element ULA with normalized wavelength. A correlated interference is present, with a 20º of elevation separation SIR (Signal-to-Interference Ratio) = 10 dB and correlation coefficient $\rho = 0.3e^{j(60\times\pi/180)}$. Beacon signals are white Gaussian processes with 1024 samples. Finally, a 50 iteration Monte Carlo simulation is performed.

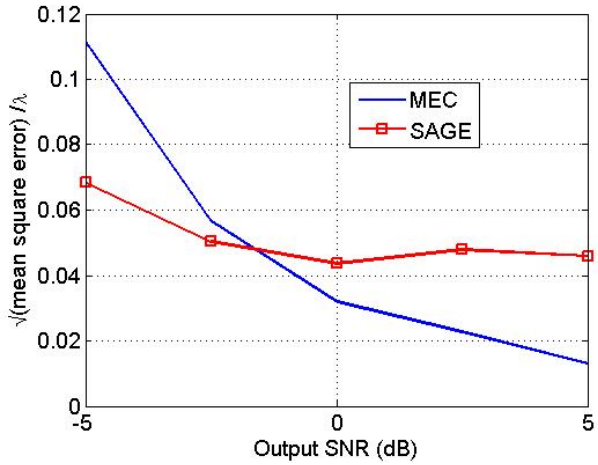

Figure 2 – MEC and SAGE mean square error

In the preceding figure a ±0.4λ uniform position error was introduced. As can be seen, SAGE is more robust in a low SNR scenario. Its mean square error could be further decreased if more optimizations cycles were applied. However, more computing time is obviously required in this case. Positions coordinates have been decoupled in three separate problems for a simpler and more direct comparison between algorithms. Regarding variance and the CRLB:

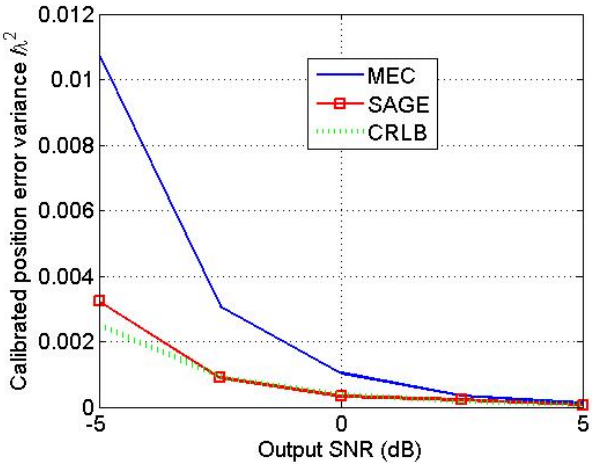

Figure 3 – MEC and SAGE variance, compared with the CRLB

MEC is a simpler algorithm than SAGE and, thus faster, with a good enough performance above -2 dB of output SNR, where residual positions errors are below 5% of the wavelength. Both are asymptotically efficient in the CRLB sense, and it can be proved that they are unbiased. The main advantage of SAGE over MEC is that it can perform a global calibration, including unknown beacon DoAs, provided that restrictions in [10] are met. However, since both are based on non-linear optimizations, careful notice must be taken to avoid solutions' ambiguities. In order to increase the convergence space, decoupling between unknowns can be performed with appropriate beacons orientations, for example.

6. CONCLUSIONS

Adaptive arrays offer interesting properties such as optimum beamforming or interference rejection. However, they must cope with calibrating errors in order to avoid serious degradations in performance. Two algorithms have been treated. SAGE is based on a MLE criterion. MEC, on the other hand, uses Burg procedure in order to perform calibration in a new way. Future simulations will seek improved non-linear optimizations methods in order to lay special emphasis on real-time behavior for on-line calibration applications.

REFERENCES

[1] A. Torre, J. Gonzalo, R. M. Pulido, R. Martínez, "New Generation Ground Segment Architecture for LEO Satellites," *in Proc. 57th IAC (International Astronautical Congress), Valencia, Spain*, 2nd-6th October, 2006.

[2] H. L. Van Trees, *Optimum Array Processing. Part IV of Detection, Estimation, and Modulation Theory*, 1st ed., Wiley, 2002.

[3] L. C. Godara, *Smart Antennas,* 1st ed., CRC Press, 2004.

[4] N. Tyler, B. Allen, and H. Aghvami, "Adaptive Antennas: The Calibration Problem", *Adaptive Antennas and MIMO Systems for Wireless Systems, IEEE Communications Magazine, pp. 114-122*, December 2004.

[5] A. J. Weiss and B. Friedlander, "Array Shape Calibration Using Sources in Unknown Locations: a Maximum Likelihood Approach", *IEEE Trans. Acoust., Speech, Signal Process., vol. ASSP-37, pp. 1958-1966*, December 1989.

[6] A. P. Dempster, N. M. Laird, D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm", *Journal of the Royal Statistical Society, Series B, Vol. 39, No.1, pp. 1-38*, 1977.

[7] J. A. Fessler and A. O. Hero, "Space-Alternating Generalized Expectation-Maximization Algorithm", *IEEE Transactions on Signal Processing, vol. 42, no. 10, pp. 2664-77*, October 1994.

[8] P. J. Chung, and J. F. Böhme, "Comparative Convergence Analysis of EM and SAGE Algorithms in DoA Estimation", *IEEE Transactions on Signal Processing, Vol. 49, No. 12*, December 2001.

[9] P. J. Chung, and S. Wan, "Array Self-Calibration Using SAGE Algorithm", *IEEE 5th Sensor Array and Multichannel Signal Processing Workshop*, July 2008.

[10] Y. Rockah and P. M. Schultheiss, "Array Shape Calibration Using Sources in Unknown Locations – Part I: Far-Field Sources", *IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-35, No. 3*, March 1987.

[11] T. M. Cover, and J. A. Thomas, *Elements of Information Theory*, 2nd ed., Wiley-Interscience, 2006.

[12] D. A. Gray, "Maximum Entropy Estimates of the Wavenumber Power Spectrum of Acoustic Data from a Linear Array of Equispaced Sensors", *Technical Report WSRL-0196-TR, Department of Defence*.

[13] M. Nagatsuka, et al., "Adaptive Array Antenna Based on Spatial Spectral Estimation using Maximum Entropy Method", *IEICE Trans. Commn*., 1994.

[14] A. J. Weiss, and B. Friedlander, "On the Cramér-Rao Bound for Direction Finding of Correlated Signals", *IEEE Trans. Signal Process*., vol. SP-41, pp. 495-499, January 1993.

APPENDIX 1: CRAMÉR-RAO LOWER BOUND

The classical CRLB has been used in simulations, considering that antennae positions can have an associated PDF but their true value does not change from one experiment to another. The same can be said about amplitude and phase imbalances if their drift is slow enough over time. Beacon DoAs are supposed to be known. In that case, there are 3N-3 desired parameters for the N-1 uncalibrated antennae, and D^2+1 parasitic ones, the Gaussian noise variance and the parameters describing the spectral density matrix of the D incoming beacon signals. The desired parameters $\boldsymbol{\theta}_{\mathrm{w}}$' Cramér-Rao matrix can be calculated as:

$$\boldsymbol{C}_{CR}(\boldsymbol{\theta}_{\mathrm{w}}) = [\boldsymbol{J}_{ww} - \boldsymbol{J}_{wu}\boldsymbol{J}_{uu}^{-1}\boldsymbol{J}_{uw}]^{-1} \tag{13}$$

$\boldsymbol{J}$ is the overall FIM (Fisher Information Matrix), $\boldsymbol{J}_{ww}$ the FIM associated with the desired parameters, $\boldsymbol{J}_{uu}$ the one related to the parasitic parameters, and the other two describe crossed couplings. Equation (13) is true if all beacon signals are present simultaneously. Should this not be the case, for example in decoupled non linear optimization, the true matrix would be:

$$\boldsymbol{C}_{CR} = [\mathrm{K}_1\boldsymbol{J}_1 + \mathrm{K}_2\boldsymbol{J}_2 + \mathrm{K}_3\boldsymbol{J}_3]^{-1} \tag{14}$$

Where K_{i} is ith signal's block size and $\boldsymbol{J}_{\mathrm{i}}$ its associated FIM. In [14] a similar derivation is performed, considering correlated interferences, although the desired parameters are the signals' DoAs.

Taking one signal block and omitting sub-index notation, the FIM becomes:

$$\boldsymbol{J}_{mn} = K\, tr(\mathbf{R}^{-1}\,\partial\mathbf{R}/\,\partial\theta_{\mathrm{m}}\mathbf{R}^{-1}\,\partial\mathbf{R}/\,\partial\theta_{\mathrm{n}}) \tag{15}$$

Where $\theta_{\mathrm{m,n}}$ are model's parameters and $tr(\ \ )$ is the trace operator. We assume that:

$$\boldsymbol{R} = \boldsymbol{V}\boldsymbol{S}_f\boldsymbol{V}^H + \sigma_n^2\,\boldsymbol{I} \tag{16}$$

Where $\boldsymbol{V}$ is the array manifold matrix, the columns of which are the input signals' **SV**, $\boldsymbol{S}_f$ the signals' spectrum matrix, σ_n^2 the noise power, and $\boldsymbol{I}$ the identity matrix. Let's start with the desired positions parameters. Consider we want the FIM coefficient corresponding to antennae l_1 and l_2's c_1 and c_2 Cartesian coordinates. This means we are looking for $(\boldsymbol{J}_{ww})_{\mathrm{m}_1\mathrm{m}_2}$ so that:

$$\mathrm{l}_{\mathrm{i}} = \left\lceil \frac{m_i+3}{3} \right\rceil \tag{17}$$

$$\mathrm{c}_{\mathrm{i}} = mod(m_i - 1, 3) + 1 \tag{18}$$

Where it is implied that position coordinates belonging to the same antenna are grouped together. Now, suppose we reorganize the desired parameter vector and calculate the FIM coefficients for each Cartesian coordinate separately. Then:

$$\boldsymbol{J}_{ww}^{\mathrm{c}_1\mathrm{c}_2} = 2\mathrm{KRe}\left\{\left[\mathbf{D}_{\mathrm{c}_2}\boldsymbol{S}_f\boldsymbol{V}^H\mathbf{R}^{-1}\right]\odot\left[\mathbf{D}_{\mathrm{c}_1}\boldsymbol{S}_f\boldsymbol{V}^H\mathbf{R}^{-1}\right]^{\mathrm{T}} + \cdots \right.$$
$$\left. \ldots + \left[\mathbf{D}_{\mathrm{c}_2}\boldsymbol{S}_f\boldsymbol{V}^H\mathbf{R}^{-1}\mathbf{V}\boldsymbol{S}_f\boldsymbol{D}_{\mathrm{c}_1}^H\right]\odot\mathbf{R}^{-\mathrm{T}}\right\} \tag{19}$$

$(\ \ )^T$ is the matrix transpose operator, Re{ } the real part operator, $\odot$ the Hadamard product and:

$$\mathbf{D}_{\mathrm{c_i}} = \mathbf{V}\odot\begin{pmatrix} \mathrm{j}\boldsymbol{k}_1(\mathrm{c_i}) & \cdots & \mathrm{j}\boldsymbol{k}_D(\mathrm{c_i}) \\ \vdots & \cdots & \vdots \\ \mathrm{j}\boldsymbol{k}_1(\mathrm{c_i}) & \cdots & \mathrm{j}\boldsymbol{k}_D(\mathrm{c_i}) \end{pmatrix} \tag{20}$$

$\boldsymbol{k}_D(\mathrm{c_i})$ is the dth incoming signal wavenumber vector, as defined in (12).

Regarding unwanted parameters, we will treat separately the $\boldsymbol{S}_f$ coefficients and the noise power. The FIM describing coupling between desired parameters and noise variance is:

$$\boldsymbol{J}^{\mathrm{c_i}}_{w\sigma_n^2} = 2\mathrm{K}\,\mathrm{diag}\left[\mathrm{Re}\left(\mathbf{D}_{\mathrm{c_i}}\boldsymbol{S}_f\boldsymbol{V}^H\mathbf{R}^{-2}\right)\right] \tag{21}$$

diag( ) takes the main diagonal of a matrix operand. On the other hand, $\boldsymbol{J}^{\mathrm{c_i}}_{ws_f}$ calculus involves very cumbersome operations and will not be detailed in this paper. The other FIMs' description can be found in [14].

APPENDIX 2: MEC MATHEMATICAL ANALYSIS

A brief mathematical analysis of the MEC algorithm introduced in this paper will be performed. A more detailed evaluation, including statistical considerations, will be presented in another paper. This study stems from equations (9) and (11), considering one pair of antennae, a reference one and an uncalibrated one, noted k. Then

$$a_1 = -\frac{R_{k1}}{R_{kk}} \tag{22}$$

Where the corresponding correlation coefficients are used. The optimization algorithm seeks the spatial spectrum's maximum. We define:

$$F_{\boldsymbol{p}} \triangleq \left(\frac{S(\boldsymbol{p})}{\sigma^2}\right)^{-1} \tag{23}$$

And:

$$\boldsymbol{R} = \sigma_s^2\boldsymbol{So}\,\boldsymbol{So}^H + \sigma_n^2\,\boldsymbol{I} \tag{24}$$

Where σ_s^2 is the desired signal's power. Hence:

$$R_{k1} = \sigma_s^2\, e^{j\boldsymbol{k}_o^T\boldsymbol{p}_{rk}} \tag{25}$$

$$R_{kk} = \sigma_s^2 + \sigma_n^2 \tag{26}$$

Where $\boldsymbol{p}_{rk}$ are the true kth-antenna position coordinates. Taking $F_{\boldsymbol{p}}$'s gradient and equaling it to 0:

$$\begin{aligned} &\nabla_{\boldsymbol{p}}F_{\boldsymbol{p}} = -2Re\left(a_1/|a_1|e^{-j\boldsymbol{k}_o^T\boldsymbol{p}}\,\mathrm{j}\mathbf{k}_{\mathrm{o}}\right) = 0 \\ \Leftrightarrow\ &\sin(\boldsymbol{k}_o^T\boldsymbol{p}_r)\cos(\boldsymbol{k}_o^T\boldsymbol{p}) - \cos(\boldsymbol{k}_o^T\boldsymbol{p}_r)\sin(\boldsymbol{k}_o^T\boldsymbol{p}) = 0 \\ \Leftrightarrow\ &\boldsymbol{k}_o^T(\boldsymbol{p}-\boldsymbol{p}_r) = 2n\pi,\ n\in\mathbb{Z} \end{aligned} \tag{27}$$

The equation, where the sub-index k has been dropped, will hold for $\boldsymbol{p} = \boldsymbol{p}_r$. However, there is a phase ambiguity that will strongly depend on the desired signal wave vector. For decoupled Cartesians components, the correct solution will be reached when the true positions' deviations are as follows:

$$(abs(\Delta\boldsymbol{p}))_k < \lambda \tag{28}$$

If components are coupled, the equation becomes:

$$\boldsymbol{u}^T abs(\Delta\boldsymbol{p}) \ < \lambda \tag{29}$$

u is the direction cosines vector. A practical iterative optimization algorithm will have worse performance, even if no ambiguity is present, if the spectrum peak broadens. An heuristic view of this degradation can be obtained from equation (11) for $\boldsymbol{p} = \boldsymbol{p}_r$:

$$F_{\boldsymbol{p}} = \frac{\sigma^2}{\left|1-\frac{\sigma_s^2}{\sigma_s^2+\sigma_n^2}\right|^2} = \frac{\sigma^2}{|1-(1+snr^{-1})^{-1}|^2} \tag{30}$$

The optimum case is for an infinite SNR. If an interference is added, we get:

$$a_{1_{\mathrm{k}}} = -\frac{\sigma_s^2}{\sigma_s^2+\sigma_i^2+\sigma_n^2}e^{j\boldsymbol{k}_o^T\boldsymbol{p}_k} - \frac{\sigma_i^2}{\sigma_s^2+\sigma_i^2+\sigma_n^2}e^{j\boldsymbol{k}_i^T\boldsymbol{p}_k} \tag{31}$$

The new gradient equation is:

$$\sin\left(\boldsymbol{k}_o^T(\boldsymbol{p}-\boldsymbol{p}_r)\right) + \frac{\sin(\boldsymbol{k}_o^T\boldsymbol{p}-\boldsymbol{k}_i^T\boldsymbol{p}_r)}{sir} = 0 \tag{32}$$

SIR is in natural units in (32). The spurious term will tend to 0 either if the desired signal is much stronger than the interference or if both have similar DoAs, that is, $\boldsymbol{k}_o \approx \boldsymbol{k}_i$.

If the interference's power is much greater than the desired beacon signals, the estimation yields:

$$\hat{\boldsymbol{p}} \approx \left(\boldsymbol{K}_o{}^H\boldsymbol{K}_o\right)^{-1}\boldsymbol{K}_o{}^H\boldsymbol{K}_i\boldsymbol{p}_r + (2n\pi \quad 2n\pi \quad 2n\pi)^T \tag{33}$$

Where:

$$\boldsymbol{K}_o = \begin{pmatrix} \boldsymbol{k}_{o1}^T \\ \vdots \\ \boldsymbol{k}_{oM}^T \end{pmatrix} \tag{34}$$

$$\boldsymbol{K}_i = \begin{pmatrix} \boldsymbol{k}_{i1}^T \\ \vdots \\ \boldsymbol{k}_{iM}^T \end{pmatrix} \tag{35}$$

Where M non simultaneous beacon signals are used, and the interferer is supposed to move from experiment to experiment. Omitting ambiguity considerations, the estimator bias is:

$$\Delta\boldsymbol{p} = \hat{\boldsymbol{p}} - \boldsymbol{p}_r = \left(\boldsymbol{K}_o{}^{-1}\boldsymbol{K}_i - \boldsymbol{I}\right)\boldsymbol{p}_r \tag{36}$$

It will decrease for similar interferer and desired signals' power, being approximately:

$$\Delta\boldsymbol{p} = \hat{\boldsymbol{p}} - \boldsymbol{p}_r = \frac{\left(\boldsymbol{K}_o{}^{-1}\boldsymbol{K}_i - \boldsymbol{I}\right)}{\mathbf{2}}\boldsymbol{p}_r \tag{37}$$

As already implied, the preceding derivations considered a true correlation matrix instead of the complex Wishart-distributed sample one.